BioMol Concepts, Vol. 3 (2012), pp. 395–402 • Copyright © by Walter de Gruyter • Berlin • Boston. DOI 10.1515/bmc-2012-0022

Review

Tet proteins: on track towards DNA demethylation?

Nathalie Véron

Friedrich Miescher Institute for Biomedical Research, Maulbeerstrasse 66, CH-4058 Basel, Switzerland

e-mail: nathalie.veron@fmi.ch

Abstract

Dynamic DNA methylation is a prerequisite for many developmental processes and maintenance of cellular integrity. In mammals however, mechanisms of active DNA demethylation have for long been elusive. The discovery of the ten-eleven translocation (Tet) family of enzymes that oxidize 5-methylcytosine (5mC) to 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC) or 5-carboxylcytosine (5caC) provided new means by which DNA methylation could actively be reversed. This review focuses on the possible mechanisms of DNA demethylation via Tet proteins and their metabolites 5hmC, 5fC and 5caC. Additionally, it discusses the roles of the three Tet protein family members Tet1, Tet2 and Tet3 as developmental regulators, probably in part independent of their enzymatic activity. By contrast, recent evidence suggests a function of 5hmC as an epigenetic mark on its own, going beyond the expectation of only acting as an intermediate in an active DNA demethylation pathway.

Keywords: DNA demethylation; epigenetics; 5-hydroxymethylcytosine; 5-methylcytosine; Tet proteins.

Introduction

Development of a healthy organism requires the orchestration of complex mechanisms for correct interpretation of the underlying genomic information. This process starts with the formation of a totipotent one-cell embryo that needs to configure its gene expression program to ensure proper subsequent cell divisions and differentiation of tissues and organs forming an adult organism.

Cellular differentiation involves transcription factor mediated gene regulation and flexible modification of chromatin structure (DNA tightly wrapped around basic histones forming nucleosomes). Combinations of various post-translational modifications of histones (e.g., methylation or acetylation), as well as addition of small chemical groups (e.g., methyl groups) to DNA bases constitute the heritable epigenetic profile of a specific cell type. As a consequence, a cell-intrinsic gene expression program is established and maintained.

DNA methylation, here referred to as the methylation of the cytosine base at the fifth position of the pyrimidine ring (5-methylcytosine, 5mC) occurs almost exclusively in a CpG context (where a guanine follows a cytosine in the DNA sequence) (1). DNA methylation has been associated with a variety of different regulatory processes in mammals. It is a hallmark of differentiated cells, where it mediates stable gene silencing that is essential for the maintenance of cellular integrity (2, 3). DNA methylation is also involved in the control of gene dosage in the context of X chromosome inactivation (the process whereby one of the two X chromosomes in females is transcriptionally silenced) and genomic imprinting, where it controls the parent-of-origin expression of alleles via the sex-specific methylation of imprint control regions (4–6). It is furthermore needed to repress transposable elements, whose activation might otherwise damage the DNA (7). Methylation of cytosine on DNA is catalyzed by the DNA methyltransferases (Dnmts) Dnmt1, Dnmt 3a and 3b, which are required for maintenance and *de novo* methylation, respectively (8–10).

Although DNA methylation is essential for survival, some natural processes require its reversal (11). Upon fertilization of the oocyte, the paternal (sperm) genome rapidly loses global methylation (12). Later in embryogenesis, primordial germ cells (PGCs), the developing 'founder' germ cells that are formed in implanted embryos, erase their parentally inherited genomic imprints. After erasure, PGCs acquire new imprints according to their sex. These are then ultimately transmitted to the subsequent generation by means of male and female germ cells in the adult organism (13). Interestingly, loss of DNA methylation often coincides with the requirement of reaching a pluripotent cell fate, such as in PGCs and in cells of the early embryo. Although in mammals scenarios involving rapid decrease in DNA methylation have been known for years, a longstanding search for pathways involved has only recently begun to reveal some of the details (14). By contrast, active DNA demethylation pathways in plants have been extensively studied. Plants harbor enzymes of the base excision repair (BER) machinery that recognize 5mC as a substrate (15). The modified base is subsequently removed by the activity of a glycosylase that cuts the sugar backbone of the DNA and replaces the gap by the unmodified base analog. However, animals lack homologous BER enzymes having high affinity towards 5mC (14, 15). One widely accepted scenario for removing DNA methylation is the inhibition of Dnmts, whose function in the maintenance of methylation occurs through recognition of hemimethylated DNA and propagation of 5mC upon replication (16–18). However, this passive loss of methylation is slow and requires cell division.

Therefore, it cannot explain the rapid demethylation, for example, of the paternal genome in the fertilized oocyte in the absence of cell division. Until recently, it had remained unclear what mechanisms might trigger active DNA demethylation, as none of the proposed models had so far provided sufficient answers (14, 15, 17, 18).

A new enzyme class converting methylated cytosines

The discovery of a conserved enzyme family generating the 'sixth base' 5-hydroxymethylcytosine (5hmC) focused broad attention. The ten-eleven translocation (Tet) protein family of 2-oxoglutarate (2OG)- and Fe(II)-dependent enzymes can convert 5mC to 5hmC by hydroxylation of the methyl group (19, 20). At the same time, Kriaucionis and Heintz reported abundant 5hmC, originally identified as nucleic acid component in viruses, in mammalian brain cells (21, 22). The detection of Tet proteins and 5hmC in mammals provided a new direction in the search for active DNA demethylation pathways, as 5hmC could serve as an intermediate to generate unmodified cytosine.

The Tet enzyme family consists of three members, Tet1, 2 and 3, which have a distinct expression pattern throughout development but tend to be expressed highest in undifferentiated, pluripotent cell types (Figure 1). Tet3 levels are highest in oocytes persisting after fertilization and decreasing with the first two embryonic cleavages, when Tet1 and 2 become activated (23). Whereas Tet1 and, to a lesser extent, Tet2 become abundant in preimplantation embryos, their protein levels decrease in cells undergoing differentiation in implanted developing embryos with the exception of some cell types of the brain (23). In adults, Tet2 is highly expressed in cells of the hematopoietic system and is frequently downregulated or subjected to loss-of-function mutations in related cancer cells, which is reviewed in (24). Concomitant with the presence of Tet family proteins, 5hmC is relatively abundant in oocytes, in the early embryo and in some brain cells (25–27). By contrast, most mammalian adult tissues show very low levels of 5hmC (25, 27). Hence, the presence of Tet proteins and 5hmC globally anticorrelates with cellular differentiation and high DNA methylation levels.

Tet protein function in mouse embryonic stem cells

Pluripotent embryonic stem cells (ESCs) are derived from the inner cell mass (ICM), a cluster of pluripotent cells that will form the embryo proper after implantation of the blastocyst (a late preimplantation embryo bearing an outer layer of epithelial trophoblast cells (trophectoderm, TE) surrounding a fluid-filled cavity and the ICM). ESCs express high levels of Tet1 and intermediate levels of Tet2 and, consistently, abundant 5hmC (20, 28–32). Hence, these pluripotent cells have served as an excellent model system to study biological functions of Tet proteins and DNA demethylation mechanisms via 5hmC.

One of the first descriptions of Tet1 function focused on its role in the maintenance of pluripotency in ESCs (20). Lentivirus-mediated depletion of Tet1 protein by knockdown (kd) approaches using small hairpin RNAs (shRNAs) impaired the capacity of ESC self-renewal and led to upregulation of differentiation markers. The authors of this study describe a

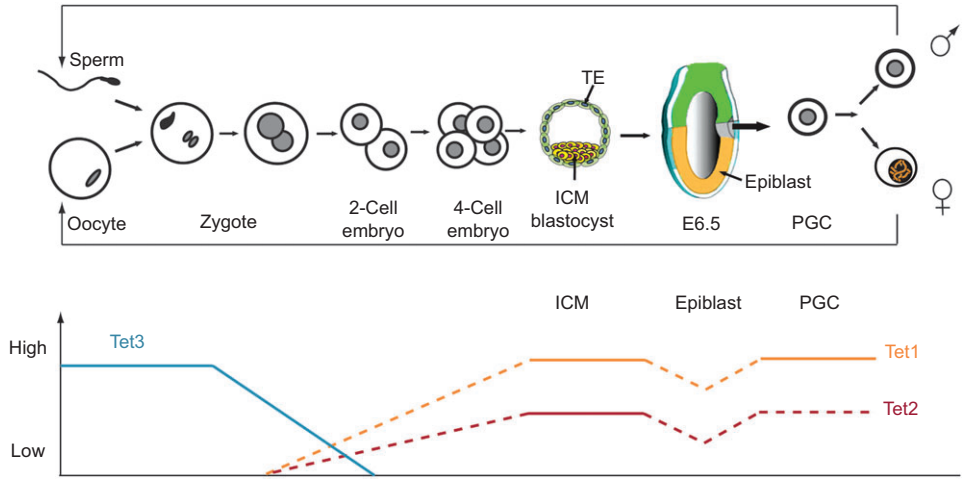

Figure 1 Tet protein dynamics in early embryonic development.
Tet3 is abundant in the oocyte and remains present at high levels after fertilization in the zygote. In two-cell embryos Tet3 levels rapidly decrease, when *Tet1* and *Tet2* start to become activated. Tet1 and Tet2 are highly expressed in the inner cell mass (ICM) but not in the trophectoderm (TE) of the blastocyst (the late preimplantation embryo). Tet protein levels decrease upon differentiation of embryonic stem cells [(ESCs) derived from the ICM) and, therefore, probably also in the epiblast (cells of the implanted embryo, which will give rise to the embryo proper]. At embryonic day (E) 6.5 primordial germ cells (PGCs), which will differentiate into mature germ cells, are specified from the posterior proximal epiblast. Tet1 and Tet2 levels increase upon PGC development. Dashed lines represent putative and straight lines represent experimentally validated expression.

crosstalk between the pluripotency factor *Nanog* and Tet1 by its binding to the *Nanog* promoter, which probably causes the altered ESC fate upon Tet1 removal. Nevertheless, the function of Tet1 in maintenance of pluripotency in ESCs is still debated (20, 29, 30, 33, 34). Several independent studies of genome-wide Tet1 occupancy in ESCs using different antibodies against Tet1 have shown a high degree of overlap of Tet1 binding sites, to as much as 90% between comparable analyses (30–32, 35). These studies revealed that Tet1 prefers binding within intragenic regions, predominantly at regions that contain clusters of CpG dinucleotides forming CpG islands (CGIs) present in two-thirds of all vertebrate gene promoters. These regions are generally hypomethylated in ESCs and enriched for histone modification Histone 3 lysine 4 trimethylation (H3K4me3), an epigenetic mark associated with active transcription (2, 36, 37). At some CGIs additional Histone 3 lysine 27 trimethylation (H3K27me3), a mark indicating gene repression, is present generating a bivalent epigenetic state (2). Tet1 is also enriched around transcriptional start sites (TSSs) of CpG rich promoters, suggesting a role in transcriptional regulation of these genes (30–32, 35). Indeed, depletion of Tet1 by kd approach alters the gene expression level of some of its targets. Although the predicted function of Tet1 was to counteract repressive DNA methylation, not all Tet1 target genes become silenced upon removal of the enzyme. In fact, only 10% of Tet1 marked genes alter their expression level and more than half of the latter are upregulated upon Tet1 kd, suggesting that Tet1 can play a dual role in transcriptional regulation (30–32, 35). Thus, depletion of the Tet1 enzyme leads to a global reduction of 5hmC reflecting its role in enzymatic conversion of 5mC in ESCs. Simultaneously, only a minor global, but a detectable local increase in DNA methylation at single genes becomes apparent, which might contribute to transcriptional regulation at these sites. Tet1-activated genes are enriched for those having a housekeeping function, whereas the Tet1-repressed fraction mainly harbors developmental regulators. Co-occurrence of Tet1 with either H3K4me3 alone or with bivalent H3K4me3 and H3K27me3 modifications seems to have a predictive potential of active transcription or gene silencing, respectively (30, 31). Polycomb repressive complex 2 (PRC2) proteins that establish repressive H3K27me3 marks overlap with around 95% of all Tet1-repressed genes (30, 31). Of these a major fraction loses PRC2 occupancy upon Tet1 depletion in ESCs, although no direct interaction of Tet1 and PRC2 complex members was identified. This suggests that PRC2 recruitment to most of its target genes depends on Tet1 binding, although the mechanism of this interplay remains elusive. A subfraction of the Tet1-repressed genes shows colocalization with the Sin3A co-repressor, which directly binds to Tet1 and presumably mediates transcriptional silencing through interaction with chromatin remodeling complexes at these sites (30).

The function of Tet1 in regulating gene expression might be in part independent of its enzymatic activity. Interestingly, the lack of DNA methylation (and hydroxymethylation) in ESCs (induced by mutation of all three Dnmt enzymes [Dnmt triple knockout (TKO) ESCs]) does not affect their self-renewal and pluripotency (38). These cells even retain a gene expression pattern similar to that of wild type ESCs and only show cellular defects upon induction of differentiation (39). TKO ESCs additionally depleted for Tet1 by kd show a similar set of up- and downregulated genes as compared to wild type ESCs upon Tet1 kd. This result suggests a regulatory role for Tet1 in transcription beyond its catalytic activity, as 5hmC is absent in TKO ESCs (30). However, this requires confirmation, for example, through genetic approaches using catalytically inactive Tet proteins.

Establishment of techniques for genome-wide analysis of 5hmC

As 5hmC had been a largely uncharacterized modified base, its functional analysis in the mammalian genome provided a technical challenge. The classical technique to distinguish methylated from unmethylated cytosines, bisulfite conversion of DNA, is not suitable for the distinction of 5mC and 5hmC bases (40). As commonly used methods to investigate the role of 5hmC were unavailable, many groups developed specific techniques for mapping genome-wide distribution of the hydroxylated base. Several 5hmC specific antibodies were generated and were applied for hydroxymethyl-DNA immunoprecipitation followed by high-throughput sequencing (hMeDIP-Seq) or to whole-genome tiling microarrays (23, 28–30, 32). In addition, several chemical labeling and enzyme-based methods for detection of 5hmC in the genome were established (28, 41). Two chemical labeling based approaches are supposed to be highly sensitive to detection of even single 5hmC bases (28). One method entails enzymatic addition of one glucose molecule to the hydroxymethyl group in 5hmC (glucose ligation biotinylation, GLIB). The covalently linked glucose is subsequently oxidized and coupled to biotin, which is then isolated by its affinity to streptavidin. Another conversion based method takes advantage of chemical conversion of 5hmC to cytosine 5-methylenesulfonate (CMS), which is recognized by an anti-CMS antibody. The GLIB and the CMS method were coupled to massive parallel sequencing for genome-wide localization of 5hmC in ESCs, providing similar results but increased sensitivity as compared to the use of anti-5hmC antibodies. However, relative genome-wide comparison of 5mC and 5hmC levels had not been possible until very recently: the oxidative bisulfite sequencing method for single base resolution allows to quantify levels of 5hmC relative to 5mC in the genome (42).

Functional roles of 5 hmC: demethylation intermediate or regulatory mark?

Similar to Tet1, 5hmC is enriched at intragenic regions and enhancers in ESCs (23, 28, 30, 43). However, in contrast to the enzyme Tet1, 5hmC seems to be preferentially present at genes with intermediate CpG content and to be less abundant at CGIs, revealing some non-overlapping sites of Tet1 enzyme and its product in the genome (30, 42, 43). The Tet1 non-bound sites positive for 5hmC could reflect the contribution of

Tet2 to hydroxymethylation at so far unknown Tet2 target sites in ESCs, as it is also expressed in ESCs but has not yet been mapped in the genome. Alternatively, 5hmC might function as a regulatory mark on its own, for example, regulating transcription, serving for repulsion of Dnmts or regulating accessibility of epigenetic players such as chromatin remodelers.

Some CGIs show low 5hmC but abundant Tet1 binding arguing for rapid processing of 5hmC at these sites. One suggested pathway includes the action of AID/Apobec (activation-induced cytidine deaminase/apolipoprotein B mRNA-editing, enzyme catalytic, polypeptide family) deaminases that convert 5hmC to 5-hydroxymethyluracil (5hmU) (44). The latter is subsequently processed by glycosylases that cut the DNA backbone, for example, Thymine DNA glycosylase (TDG) or single-strand monofunctional uracil DNA glycosylase (SMUG1) (44, 45). An alternative model proposes that Tet proteins oxidize 5mC to 5hmC that is subsequently oxidized to 5-formylcytosine (5fC) and finally to 5-carboxylcytosine (5caC) (46). The oxidation product 5caC serves as a substrate for the glycosylase TDG and is replaced by unmodified cytosine (47). Alternatively, 5caC could be decarboxylated to generate unmodified cytosine, although no enzyme for this step has been identified so far (46). The process of full oxidation of 5mC to 5caC might be dependent on the stability of Tet enzyme binding, which seems to be increased at CpG dense regions. It is also conceivable that unknown co-factors add to this process and that several pathways contribute additively to DNA demethylation, possibly in a chromatin structure or cell type-dependent manner.

The high abundance of 5hmC and the reduced presence of Tet1 at sites of intermediate CpG density suggest a transient low-affinity binding of the enzyme and subsequent stabilization of hydroxymethylation in this context (42). The stability of 5hmC suggests a function independent of being an active demethylation intermediate. However, presence of DNA hydroxymethylation within genes does not correlate with either an active or a repressed transcriptional state, suggesting that 5hmC is not a predictive mark for transcription *per se* (30, 48). Distribution of 5hmC along the length of genes revealed distinct density profiles that seem to be more predictive for gene activity: genes that are depleted from 5hmC around TSSs, but contain higher 5hmC levels at their 3 ends tend to be active. By contrast, genes harboring abundant 5hmC at TSSs and promoter sequences tend to be silent, although results from different studies are somewhat inconsistent (29, 35, 48).

Unlike Tet1, 5hmC is enriched at enhancer regions, which regulate the transcriptional activity of linked genes. Interestingly, these regions typically show relatively low methylation levels in ESCs (42, 49, 50). This might reflect sequence-specific binding of transcription factors contributing to an open chromatin configuration, as methylated cytosines recruit methyl-CpG-binding domain proteins (Mbds) that support gene silencing (50, 51). Hence, conversion of 5mC to 5hmC might inhibit the binding of repressive Mbds or of Dnmts, thereby generating a hypomethylated state accessible for regulatory complexes for induction of transcriptional activity at downstream genes.

A recent study that aimed at investigating the role of several chromatin remodelers in ESC pluripotency revealed a new player in the read-out of 5hmC. Unlike other Mbd family members that recognize 5mC only, the methyl-binding domain protein 3 (Mbd3) binds to hydroxymethylated DNA, probably due to the unique structure of its Mbd binding pocket (52). Mbd3 is a subunit of the silencing Nucleosome Remodeling and Deacetylase (NuRD) chromatin remodeling complex and colocalizes with PRC2 and Tet1 at sites possessing elevated 5hmC levels in ESCs. Moreover, depletion of Mbd3 by a kd approach in ESCs leads to a loss of 5hmC and impaired Tet1 binding, suggesting a role for Mbd3 in the maintenance of hydroxymethylated states. By contrast, Tet1 is needed to recruit Mbd3, suggesting a positive feedback loop of the two interactors. Interestingly, Brg1, a member of the activating Swi/Snf chromatin remodeling complex, which antagonizes NuRD, also binds to sites occupied by Mbd3. The presence of two antagonistic complexes at a common set of targets might serve to keep an intact balance between repressive and active transcriptional domains creating a platform for transcriptional plasticity in pluripotent cells (52). However, it still remains unclear if 5hmC interacts similarly with chromatin remodeling complexes in other cell types.

Tet3 function in oocytes and preimplantation embryos

Upon fertilization of the oocyte, the paternal genome rapidly loses 5mC and gains DNA hydroxymethylation. This is probably due to the activity of maternally provided Tet3, which is abundant in the oocyte, although the maternal genome itself shows little 5hmC (Figure 1) (53, 54). Hydroxymethylation established on the paternal genome remains globally stable through the first cell divisions and shows a loss, which appears exclusively replication-dependent, as revealed by immunofluorescence stainings (26). However, this observation does not exclude the possibility that single genomic regions or genes convert 5hmC to unmethylated cytosine. Indeed, conditional *Tet3* knockout (ko) in oocytes (driven by the *Zona pellucida 3* (*Zp3*) promoter) leads to delayed activation of a paternally provided allele of the pluripotency marker *Oct4*, which is methylated in sperm and appears to be demethylated to allow for gene expression in the embryo (54). Nevertheless, it remains unclear if the global persistence of 5hmC at the paternal genome in preimplantation embryos serves a crucial role in early embryonic development. Oocyte-specific deletion of *Tet3* leads to developmental arrest of around half of embryos at mid-embryogenesis, but the other half are born alive and show normal fertility (54). As a mixed genetic background was used in these experiments, cellular responses to the lack of Tet3 might depend on the individual genetic context leading to differential penetrance of the mutant phenotype. Alternatively, *Zp3*-driven conditional *Tet3* deletion might, in some oocytes, lead to the presence of residual levels of the enzyme, because deletion may occur at slightly different time points in oogenesis. As

a consequence, low levels of Tet3 might just be enough to safeguard embryonic development. This possibility could be addressed by the use of an earlier promoter for conditional *Tet3* deletion. To date, detailed characterization of mechanisms involving Tet3 in epigenetic gene regulation remains elusive, but its genome-wide profiling in preimplantation embryos provides a technical challenge due to the small amount of cells (and DNA) available. Moreover, it is unclear if 5hmC that persists in early embryos is needed for transcriptional regulation by recruitment of chromatin remodelers as has been described for ESCs (52). Interestingly, preimplantation development of parthenotes (generated by activation of an unfertilized oocyte and exclusively bearing the maternal genome) lacking high levels of 5hmC is unimpaired, suggesting that hydroxymethylation of DNA is dispensable, at least in this specific scenario (26).

The role of Tet proteins in embryonic development

During the first cellular divisions of the fertilized oocyte Tet3 is the major Tet family protein that is active, with Tet1 and Tet2 becoming abundant only in the blastocyst (Figure 1) (20, 33, 34). So far, ESC-based functional studies of Tet1 have suggested an essential role in maintenance of the pluripotent state and in transcriptional up- and downregulation of distinct target gene sets, at least partly by generating 5hmC. However, *in vivo* analysis of Tet1 function using a *Tet1* ko approach by homologous recombination revealed that mice lacking Tet1 are viable and fertile, although smaller at birth (34). Moreover, *Tet1* ko ESCs are self-renewing and pluripotent contradicting previous findings using *Tet1* kd ESCs that lose pluripotency features and show skewing of their differentiation potential (20, 29, 33). Gene expression profiling of *Tet1* ko ESCs revealed a deregulation of around 200 genes, of which 60% were downregulated and 40% upregulated at least twofold. Compared to previous studies of Tet1 depletion using shRNA-mediated kd approaches, this only represents around one-sixth of genes previously found to be deregulated (29–31). When comparing several individual Tet1 kd studies, Tet1-deregulated genes show a rather poor overlap of around 10%, although Tet1 binding sites were around 90% identical (35). Nevertheless, all studies show relatively similar patterns of upregulated and downregulated genes, confirming the dual function of Tet1 in transcriptional regulation. The discrepancies between *Tet1* ko and kd studies might be due to technical limitations of kd approaches such as off-target effects of retroviral-based constructs and incomplete depletion of transcripts and protein. By contrast, it is conceivable that variable expression levels of Tet enzymes (e.g., by different protein depletion efficiencies) might generate different cellular outcomes indicating biological variability in response to absolute Tet protein levels. Moreover, complete loss of Tet1 by a ko approach might cause Tet2 upregulation in ESCs, whereas incomplete removal of Tet1 by kd might not do so (or only to a lesser extent).

DNA demethylation via Tet in primordial germ cells?

Despite global DNA demethylation in the zygote and maintenance of a hypomethylated state during the first embryonic cleavages, imprinted loci are unaffected and remain differentially methylated. Shortly after implantation of the mouse embryo, some posterior proximal cells of the epiblast (the cells that give rise to the embryo proper in the implanted embryo) are specified to develop PGCs at embryonic day (E) 6.25 (55, 56). Nascent PGCs migrate to the genital ridges and populate the gonads between around E10.5 and E12.5, where they will differentiate into mature germ cells. As PGCs are derived from embryonic cells that have already acquired some somatic differentiation, they need to undergo epigenetic reprogramming towards a cellular state of higher plasticity (57). Moreover, for the proper setting of sex-specific genomic imprints in germ cells, the parentally inherited differentially methylated regions (DMRs) of somatic cell origin need to be erased. Between E8 and E11.5, PGCs are extensively demethylated and consistently they also display increased expression of *Tet1* and *Tet2* around E11 (Figure 1) (58). However, to date, it is unclear to what extent Tet enzymes contribute to imprint erasure. The observed loss of 5mC might be due to conversion to 5hmC or its processing to unmodified cytosine by oxidation and/or BER (58). Although genetic approaches have shown that two players in BER, AID and TDG are at least in part required for proper demethylation in PGCs, their absence does not lead to a significant loss of PGCs (45, 59, 60). During the biggest wave of global decrease of 5mC at around E11.5, no evidence has been found so far for the presence of AID or TDG (60, 61). Thus, other BER enzymes might be involved in the processing of 5hmC in PGCs, either via metabolism to 5hmU or via Tet-mediated oxidation products 5fC and 5caC. Alternatively, multiple parallel pathways of replication-dependent passive demethylation together with Tet-mediated active demethylation have been proposed as a high-fidelity mechanism for epigenetic reprogramming in developing germ cells (58). However, this model awaits confirmation being challenged by the limited technical accessibility of PGCs.

Expert opinion

The flurry of recent studies about Tet proteins and their products 5hmC, 5fC and 5caC have clearly demonstrated their importance in developmental processes. However, to date, most of the mechanistic information has been generated in ESCs, which are a rather limited model system in light of the cellular complexity of an entire organism. Especially with respect to the differential expression of the three Tet protein family members throughout the body, comprehensive functional *in vivo* data are still missing. Moreover, it is unclear if all three Tet enzymes could play redundant roles, as they might have distinct binding sites and interact with different complexes, such as chromatin remodelers

or transcription factors in different cell types. Moreover, dependent on the interaction partners available, Tet proteins might trigger varying cellular responses as a result of their potential to regulate transcription and/or to generate DNA demethylation intermediates. The role of each player in this enzyme family needs to be further elucidated in different cell types, for example, by genome-wide analysis of their binding sites and biochemical studies of Tet protein interactors. In addition, the relationship between Tet proteins and hydroxymethylation in epigenetic regulation is still unclear. Recent studies suggest that 5hmC is more than just a DNA demethylation intermediate and that it could function as an epigenetic mark on its own recruiting chromatin remodeling complexes, as well as possibly transcription factors, histone variants or histone modifying enzymes. With respect to DNA demethylation via 5hmC, different downstream pathways seem to exist, perhaps even in parallel to ensure accuracy of the process.

Outlook

Over the past few years, knowledge of Tet proteins and 5hmC has grown exponentially owing to the development of new techniques. The constant improvement of methods for even quantitative analysis of hydroxymethylation in the genome has allowed detailed analysis that will certainly be even more refined in the future. Genetic approaches aiming at conditional deletion of more than one Tet enzyme at a time will provide insight into their (possible redundant) function in tissues and organs. The necessity of the catalytic activity of Tet for their biological function might be addressed by expressing catalytically dead enzymes resulting in a loss of 5hmC in various cellular contexts. In the future, the establishment of a comprehensive network of Tet proteins and their interaction partners, as well as its role in hydroxymethylation and DNA demethylation in the context of development and disease is to be awaited.

Highlights

- Tet proteins convert 5mC to 5hmC, 5fC and 5caC thereby contributing to active DNA demethylation in mammalian genomes.
- Tet proteins are transcriptional regulators, possibly partly independent of their catalytic activity.
- The redundant function of the three Tet protein family members is currently unclear.
- 5hmC might function as an intermediate towards DNA demethylation and as a regulatory mark.
- Tissue specificity of Tet- and 5hmC-mediated processes needs to be addressed.
- The establishment of a comprehensive epigenetic network involving Tet proteins, their metabolites, chromatin regulators and transcription factors in different cell types is to be awaited.

Acknowledgements

I would like to apologize for all related articles that have not been cited owing to space constraints. I thank Mark Gill and Angéline Eymery for critical reading of the manuscript. N.V. is supported by a European Molecular Biology (EMBO) long-term fellowship, a research fellowship granted from the Deutsche Forschungsgemeinschaft (DFG) and by the Novartis Research Foundation. I declare no conflict of financial or other interests.

References

1. Bird A. DNA methylation patterns and epigenetic memory. Genes Dev 2002; 16: 6–21.
2. Mohn F, Weber M, Rebhan M, Roloff TC, Richter J, Stadler MB, Bibel M, Schübeler D. Lineage-specific polycomb targets and de novo DNA methylation define restriction and potential of neuronal progenitors. Mol Cell 2008; 30: 755–66.
3. Weber M, Schubeler D. Genomic patterns of DNA methylation: targets and function of an epigenetic mark. Curr Opin Cell Biol 2007; 19: 273–80.
4. Heard E, Clerc P, Avner P. X-chromosome inactivation in mammals. Annu Rev Genet 1997; 31: 571–610.
5. Li E, Beard C, Jaenisch R. Role for DNA methylation in genomic imprinting. Nature 1993; 366: 362–5.
6. Reik W, Walter J. Genomic imprinting: parental influence on the genome. Nat Rev Genet 2001; 2: 21–32.
7. Walsh CP, Chaillet JR, Bestor TH. Transcription of IAP endogenous retroviruses is constrained by cytosine methylation. Nat Genet 1998; 20: 116–7.
8. Bestor T, Laudano A, Mattaliano R, Ingram V. Cloning and sequencing of a cDNA encoding DNA methyltransferase of mouse cells. The carboxyl-terminal domain of the mammalian enzymes is related to bacterial restriction methyltransferases. J Mol Biol 1988; 203: 971–83.
9. Okano M, Xie S, Li E. Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. Nat Genet 1998; 19: 219–20.
10. Gowher H, Jeltsch A. Enzymatic properties of recombinant Dnmt3a DNA methyltransferase from mouse: the enzyme modifies DNA in a non-processive manner and also methylates non-CpG (correction of non-CpA) sites. J Mol Biol 2001; 309: 1201–8.
11. Reik W, Dean W, Walter J. Epigenetic reprogramming in mammalian development. Science 2001; 293: 1089–93.
12. Oswald J, Engemann S, Lane N, Mayer W, Olek A, Fundele R, Dean W, Reik W, Walter J. Active demethylation of the paternal genome in the mouse zygote. Curr Biol 2000; 10: 475–8.
13. Shovlin TC, Durcova-Hills G, Surani A, McLaren A. Heterogeneity in imprinted methylation patterns of pluripotent embryonic germ cells derived from pre-migratory mouse germ cells. Dev Biol 2008; 313: 674–81.
14. Niehrs C. Active DNA demethylation and DNA repair. Differentiation 2009; 77: 1–11.
15. Gehring M, Reik W, Henikoff S. DNA demethylation by DNA repair. Trends Genet 2009; 25: 82–90.
16. Song J, Teplova M, Ishibe-Murakami S, Patel DJ. Structure-based mechanistic insights into DNMT1-mediated maintenance DNA methylation. Science 2012; 335: 709–12.
17. Ooi SK, Bestor TH. The colorful history of active DNA demethylation. Cell 2008; 133: 1145–8.

18. Ooi SK, Bestor TH. Cytosine methylation: remaining faithful. Curr Biol 2008; 18: R174–6.
19. Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, Agarwal S, Iyer LM, Liu DR, Aravind L, Rao A. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. Science 2009; 324: 930–5.
20. Ito S, D'Alessio AC, Taranova OV, Hong K, Sowers LC, Zhang Y. Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. Nature 2010; 466: 1129–33.
21. Wyatt GR, Cohen SS. The bases of the nucleic acids of some bacterial and animal viruses: the occurrence of 5-hydroxymethylcytosine. Biochem J 1953; 55: 774–82.
22. Kriaucionis S, Heintz N. The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. Science 2009; 324: 929–30.
23. Wu H, Zhang Y. Mechanisms and functions of Tet protein-mediated 5-methylcytosine oxidation. Genes Dev 2011; 25: 2436–52.
24. Cimmino L, Abdel-Wahab O, Levine RL, Aifantis I. TET family proteins and their role in stem cell differentiation and transformation. Cell Stem Cell 2011; 9: 193–204.
25. Globisch D, Munzel M, Muller M, Michalakis S, Wagner M, Koch S, Brückl T, Biel M, Carell T. Tissue distribution of 5-hydroxymethylcytosine and search for active demethylation intermediates. PLoS One 2010; 5: e15367.
26. Inoue A, Zhang Y. Replication-dependent loss of 5-hydroxymethylcytosine in mouse preimplantation embryos. Science 2011; 334: 194.
27. Ruzov A, Tsenkina Y, Serio A, Dudnakova T, Fletcher J, Bai Y, Chebotareva T, Pells S, Hannoun Z, Sullivan G, Chandran S, Hay DC, Bradley M, Wilmut I, De Sousa P. Lineage-specific distribution of high levels of genomic 5-hydroxymethylcytosine in mammalian development. Cell Res 2011; 21: 1332–42.
28. Pastor WA, Pape UJ, Huang Y, Henderson HR, Lister R, Ko M, McLoughlin EM, Brudno Y, Mahapatra S, Kapranov P, Tahiliani M, Daley GQ, Liu XS, Ecker JR, Milos PM, Agarwal S, Rao A. Genome-wide mapping of 5-hydroxymethylcytosine in embryonic stem cells. Nature 2011; 473: 394–7.
29. Ficz G, Branco MR, Seisenberger S, Santos F, Krueger F, Hore TA, Marques CJ, Andrews S, Reik W. Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation. Nature 2011; 473: 398–402.
30. Williams K, Christensen J, Pedersen MT, Johansen JV, Cloos PA, Rappsilber J, Helin K. TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity. Nature 2011; 473: 343–8.
31. Wu H, D'Alessio AC, Ito S, Xia K, Wang Z, Cui K, Zhao K, Sun YE, Zhang Y. Dual functions of Tet1 in transcriptional regulation in mouse embryonic stem cells. Nature 2011; 473: 389–93.
32. Xu Y, Wu F, Tan L, Kong L, Xiong L, Deng J, Barbera AJ, Zheng L, Zhang H, Huang S, Min J, Nicholson T, Chen T, Xu G, Shi Y, Zhang K, Shi YG. Genome-wide regulation of 5hmC, 5mC, and gene expression by Tet1 hydroxylase in mouse embryonic stem cells. Mol Cell 2011; 42: 451–64.
33. Koh KP, Yabuuchi A, Rao S, Huang Y, Cunniff K, Nardone J, Laiho A, Tahiliani M, Sommer CA, Mostoslavsky G, Lahesmaa R, Orkin SH, Rodig SJ, Daley GQ, Rao A. Tet1 and Tet2 regulate 5-hydroxymethylcytosine production and cell lineage specification in mouse embryonic stem cells. Cell Stem Cell 2011; 8: 200–13.
34. Dawlaty MM, Ganz K, Powell BE, Hu YC, Markoulaki S, Cheng AW, Gao Q, Kim J, Choi S-W, Page DC, Jaenisch R. Tet1 is dispensable for maintaining pluripotency and its loss is compatible with embryonic and postnatal development. Cell Stem Cell 2011; 9: 166–75.
35. Wu H, Zhang Y. Tet1 and 5-hydroxymethylation: a genome-wide view in mouse embryonic stem cells. Cell Cycle 2011; 10: 2428–36.
36. Thomson JP, Skene PJ, Selfridge J, Clouaire T, Guy J, Webb S, Kerr AR, Deaton A, Andrews R, James KD, Turner DJ, Illingworth R, Bird A. CpG islands influence chromatin structure via the CpG-binding protein Cfp1. Nature 2010; 464: 1082–6.
37. Weber M, Hellmann I, Stadler MB, Ramos L, Paabo S, Rebhan M, Schübeler D. Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. Nat Genet 2007; 39: 457–66.
38. Tsumura A, Hayakawa T, Kumaki Y, Takebayashi S, Sakaue M, Matsuoka C, Shimotohno K, Ishikawa F, Li E, Ueda HR, Nakayama J, Okano M. Maintenance of self-renewal ability of mouse embryonic stem cells in the absence of DNA methyltransferases Dnmt1, Dnmt3a and Dnmt3b. Genes Cells 2006; 11: 805–14.
39. Jackson M, Krassowska A, Gilbert N, Chevassut T, Forrester L, Ansell J, Ramsahoye B. Severe global DNA hypomethylation blocks differentiation and induces histone hyperacetylation in embryonic stem cells. Mol Cell Biol 2004; 24: 8862–71.
40. Jin SG, Wu X, Li AX, Pfeifer GP. Genomic mapping of 5-hydroxymethylcytosine in the human brain. Nucleic Acids Res 2011; 39: 5015–24.
41. Song CX, Szulwach KE, Fu Y, Dai Q, Yi C, Li X, Li Y, Chen C-H, Zhang W, Jian X, Wang J, Zhang L, Looney TJ, Zhang B, Godley LA, Hicks LM, Lahn BT, Jin P, He C. Selective chemical labeling reveals the genome-wide distribution of 5-hydroxymethylcytosine. Nat Biotechnol 2010; 29: 68–72.
42. Booth MJ, Branco MR, Ficz G, Oxley D, Krueger F, Reik W, Balasubramanian S. Quantitative sequencing of 5-methylcytosine and 5-hydroxymethylcytosine at single-base resolution. Science 2012; 336: 934–7.
43. Guo JU, Su Y, Zhong C, Ming GL, Song H. Emerging roles of TET proteins and 5-hydroxymethylcytosines in active DNA demethylation and beyond. Cell Cycle 2011; 10: 2662–8.
44. Guo JU, Su Y, Zhong C, Ming GL, Song H. Hydroxylation of 5-methylcytosine by TET1 promotes active DNA demethylation in the adult brain. Cell 2011; 145: 423–34.
45. Cortellino S, Xu J, Sannai M, Moore R, Caretti E, Cigliano A, Le Coz M, Devarajan K, Wessels A, Soprano D, Abramowitz LK, Bartolomei MS, Rambow F, Bassi MR, Bruno T, Fanciulli M, Renner C, Klein-Szanto AJ, Matsumoto Y, Kobi D, Davidson I, Alberti C, Larue L, Bellacosa A. Thymine DNA glycosylase is essential for active DNA demethylation by linked deamination-base excision repair. Cell 2011; 146: 67–79.
46. Ito S, Shen L, Dai Q, Wu SC, Collins LB, Swenberg JA, He C, Zhang Y. Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. Science 2011; 333: 1300–3.
47. He YF, Li BZ, Li Z, Liu P, Wang Y, Tang Q, Ding J, Jia Y, Chen Z, Li L, Sun Y, Li X, Dai Q, Song C-X, Zhang K, He C, Xu G-L. Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. Science 2011; 333: 1303–7.
48. Wu H, D'Alessio AC, Ito S, Wang Z, Cui K, Zhao K, Sun YE, Zhang Y. Genome-wide analysis of 5-hydroxymethylcytosine distribution reveals its dual function in transcriptional regulation in mouse embryonic stem cells. Genes Dev 2011; 25: 679–84.
49. Stadler MB, Murr R, Burger L, Ivanek R, Lienert F, Schöler A, van Nimwegen E, Wirbelauer C, Oakeley EJ, Gaidatzis D, Tiwari

VK, Schübeler D. DNA-binding factors shape the mouse methylome at distal regulatory regions. Nature 2011; 480: 490–5.
50. Bird AP, Wolffe AP. Methylation-induced repression – belts, braces, and chromatin. Cell 1999; 99: 451–4.
51. Lienert F, Wirbelauer C, Som I, Dean A, Mohn F, Schubeler D. Identification of genetic elements that autonomously determine DNA methylation states. Nat Genet 2011; 43: 1091–7.
52. Yildirim O, Li R, Hung JH, Chen PB, Dong X, Ee LS, Weng Z, Rando OJ. Mbd3/NURD complex regulates expression of 5-hydroxymethylcytosine marked genes in embryonic stem cells. Cell 2011; 147: 1498–510.
53. Wossidlo M, Nakamura T, Lepikhov K, Marques CJ, Zakhartchenko V, Boiani M, Arand J, Nakano T, Reik W, Walter J. 5-Hydroxymethylcytosine in the mammalian zygote is linked with epigenetic reprogramming. Nat Commun 2011; 2: 241.
54. Gu TP, Guo F, Yang H, Wu HP, Xu GF, Liu W, Xie ZG, Shi L, He X, Jin SG, Iqbal K, Shi YG, Deng Z, Szabó PE, Pfeifer GP, Li J, Xu GL. The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. Nature 2011; 477: 606–10.
55. Surani MA. Germ cells: the eternal link between generations. CR Biol 2007; 330: 474–8.
56. Ginsburg M, Snow MH, McLaren A. Primordial germ cells in the mouse embryo during gastrulation. Development 1990; 110: 521–8.
57. Saitou M, Barton SC, Surani MA. A molecular programme for the specification of germ cell fate in mice. Nature 2002; 418: 293–300.
58. Hackett JA, Zylicz JJ, Surani MA. Parallel mechanisms of epigenetic reprogramming in the germline. Trends Genet 2012; 28: 164–74.
59. Popp C, Dean W, Feng S, Cokus SJ, Andrews S, Pellegrini M, Jacobsen SE, Reik W. Genome-wide erasure of DNA methylation in mouse primordial germ cells is affected by AID deficiency. Nature 2010; 463: 1101–5.
60. Hajkova P, Jeffries SJ, Lee C, Miller N, Jackson SP, Surani MA. Genome-wide reprogramming in the mouse germ line entails the base excision repair pathway. Science 2010; 329: 78–82.
61. Morgan HD, Dean W, Coker HA, Reik W, Petersen-Mahrt SK. Activation-induced cytidine deaminase deaminates 5-methylcytosine in DNA and is expressed in pluripotent tissues: implications for epigenetic reprogramming. J Biol Chem 2004; 279: 52353–60.

Received May 31, 2012; accepted July 5, 2012